Appendix A. Code

1. Association

    a. Roget Association:

        i. **fastText_embedding_Generate.ipynb**: Creates word embeddings for each book in the dataset.

        ii. **Filter_Global_Embedding_Space.ipynb**: Filters and selects the words embeddings present in both the global embedding and the local embedding.

        iii. **/alignment/** : This directory contains the code to align each book embedding with respect to global embedding.

        iv. **Association_Score_Generate.ipynb**: To generate association score for 1019 concepts across all books.

        v. **Final Result - Roget - Association .ipynb**: Generates results for Roget association score.

2. Word Frequency

    a. Stopword Removal:

        i. **fullbook_stopword_removal.ipynb**: Removes stopword from each book and creates a CSV with the processed dataset.

    b. LIWC:

        i. **LIWC_Frequency_Distribution_Prepare.ipynb**: Creates a CSV containing the frequency distribution of LIWC categories for each book.

        ii. **Final Result - LIWC.ipynb**: Generates the final results with LIWC.

    c. NRC:

        i. **NRC_Frequency_Distribution_Prepare.ipynb**: Creates a CSV containing the frequency distribution of NRC categories for each book.

        ii. **Final Result - NRC.ipynb**: Generates the final results with NRC.

    d. POS:

        i. **POS_Frequency_Distribution_Prepare.ipynb**: Creates a CSV containing the frequency distribution of POS categories for each book.

        ii. **Final Result - POS.ipynb**: Generates the final results with POS.

    e. Roget Thesaurus:

        i. **Roget_Frequency_Distribution_Prepare.ipynb**: Creates a CSV containing the frequency distribution of Roget's categories for each book.

        ii. **Final Result - Roget - Frequency.ipynb**: Generates the final results with Roget

Appendix B. Dataset Links

1. Book Dataset - https://github.com/sjmaharjan/emotion_flow
2. Roget - Roget's Thesaurus
3. LIWC - https://liwc.wpengine.com/
4. NRC - NRC Emotion Lexicon